
 M E S S E N G E R (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mparch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Oct 21 16:29:45 1999; Maspar time 4.57 Seconds
 Tabular output not generated. 519,801 Million cell updates/sec

Title: >US-08-978-217-7
 Description: (1-84) from US08978217.pep
 Perfect Score: 599
 Sequence: 1 NCALBELRLVFGPLDQLHA.....ELLDGQOASPHYPGCGAG 84

Scoring table: PAM 150
 Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot37
 1:swissprot

Statistics: Mean 41.198; Variance 80.907; scale 0.509

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	98	16.4	174	1	GSPM_PSEAE GENERAL SECRETION PATH	1.00e-01
2	94	15.7	277	1	Y021_SYNY3 PUTATIVE PROTEASE SLR0	3.43e-01
3	94	15.7	364	1	YBBB_ECOLI HYPOTHETICAL 41.1 KP P	3.43e-01
4	91	15.2	958	1	SYAC_YEAST ALANYL-TRNA SYNTHETASE	8.45e-01
5	91	15.2	2870	1	NDVB_RHINE 319 KD PROTEIN.	8.45e-01
6	89	14.9	395	1	APAA_MOUSE APOLIPOPROTEIN A-IV PR	1.52e+00
7	89	14.9	916	1	SECA_TREPA PREPROTEIN TRANSLOCASE	1.52e+00
8	89	14.9	952	1	IF41_YEAST EUKARYOTIC INITIATION	2.04e+00
9	88	14.7	194	1	VGFE_ECOLI HYPOTHETICAL 21.5 KP P	2.04e+00
10	88	14.7	241	1	REP_ALCEU RIBULOSE-PHOSPHATE 3-E	2.04e+00
11	88	14.7	2418	1	SPECTRIN ALPHA CHAIN,	2.04e+00
12	86	14.4	183	1	YDJA_ECOLI HYPOTHETICAL 20.1 KP P	3.63e+00
13	86	14.4	308	1	GRP3_ARTSA GLYCINE-RICH PROTEIN G	3.63e+00
14	86	14.4	317	1	APF_PIG APOLIPOPROTEIN E PRECU	3.63e+00
15	85	14.2	202	1	SODM_MOUSE SUPEROXIDE DISMUTASE I	4.82e+00
16	85	14.2	289	1	RMLA_STRMU GLUCOSE-1-PHOSPHATE TH	4.82e+00
17	85	14.2	341	1	NU2M_ANOGA NADH-UBIQUINONE OXIDOR	4.82e+00
18	85	14.2	341	1	NU2M_ANOOU NADH-UBIQUINONE OXIDOR	4.82e+00
19	85	14.2	341	1	NU2M_LATCH NADH-UBIQUINONE OXIDOR	4.82e+00
20	85	14.2	655	1	PARE_BACSU TOPOISOMERASE IV SUBUN	4.82e+00
21	85	14.2	785	1	CUL3_SCHRO CULLIN HOMOLOGY 3 (CUL-	4.82e+00
22	85	14.2	807	1	OXYB_HUMAN OXYSTEROL-BINDING PROT	4.82e+00
23	85	14.2	809	1	OXYB_RABIT OXYSTEROL-BINDING PROT	4.82e+00

24	85	14.2	849	1	FSCI_MOUSE MAJOR FIBROUS SHEATH P	4.82e+00
25	84	14.0	187	1	OX4L_RABIT OX40 LIGAND (OX40L).	6.39e+00
26	84	14.0	274	1	NU2M_DROMA NADH-UBIQUINONE OXIDOR	6.39e+00
27	84	14.0	341	1	NU2M_DROYA NADH-UBIQUINONE OXIDOR	6.39e+00
28	84	14.0	341	1	NU2M_DROME NADH-UBIQUINONE OXIDOR	6.39e+00
29	84	14.0	351	1	COTS_BACSU SPORE COAT PROTEIN S	6.39e+00
30	84	14.0	942	1	DPO1_CHLAV DNA POLYMERASE I (EC 2	6.39e+00
31	84	14.0	1220	1	DDX8_HUMAN PROBABLE ATP-DEPENDENT	6.39e+00
32	83	13.9	101	1	YDHR_ECOLI HYPOTHETICAL 11.3 KP P	8.44e+00
33	83	13.9	249	1	EIA_ADE40 EARLY E1A 27 KD PROTEI	8.44e+00
34	83	13.9	868	1	N180_YEAST N180 PROTEIN (N1P100	8.44e+00
35	83	13.9	886	1	EP15_HUMAN EPIDERMAL GROWTH FACTO	8.44e+00
36	83	13.9	1469	1	DP27_CAEEL CHROMOSOME CONDENSATIO	8.44e+00
37	82	13.7	166	1	VG17_BPPH2 EARLY PROTEIN GP17.	1.11e+01
38	82	13.7	222	1	SODM_HUMAN SUPEROXIDE DISMUTASE I	1.11e+01
39	82	13.7	222	1	SODM_MOUSE SUPEROXIDE DISMUTASE I	1.11e+01
40	82	13.7	274	1	NU2M_DROSE NADH-UBIQUINONE OXIDOR	1.11e+01
41	82	13.7	274	1	NU2M_DROSI NADH-UBIQUINONE OXIDOR	1.11e+01
42	82	13.7	568	1	MEND_HAEIN 2-SUCCINYL-6-HYDROXY-2	1.11e+01
43	82	13.7	658	1	SPEA_ECOLI BIOSYNTHETIC ARGININE	1.11e+01
44	82	13.7	864	1	ALKS_PSEOL REGULATORY PROTEIN ALK	1.11e+01
45	82	13.7	1168	1	DDX8_SCHPO PUTATIVE PRE-MRNA SPLI	1.11e+01

ALIGNMENTS

RESULT ID	1	GSPM_PSEAE	STANDARD:	PRT:	174 AA.
AC	P25061:				
DT	01-MAY-1992 (REL. 22, CREATED)				
DT	01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)				
DT	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)				
DE	GENERAL SECRETION PATHWAY PROTEIN M.				
GN	XCPZ.				
OS	PSEUDOMONAS AERUGINOSA.				
OC	BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;				
OC	PSEUDOMONAS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE: 91092255.				
RA	FILIOUX A., BALLY M., BALL G., AKRIM M., TOMMASSEN J., LAZDUSKI A.;				
RT	"Protein secretion in gram-negative bacteria: transport across the				
RT	outer membrane involves common mechanisms in different bacteria."				
RL	EMBO J. 9:4323-4329(1990).				
CC	-1- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE				
CC	EXPORT OF PROTEINS.				
CC	-1- SUBCELLULAR LOCATION: INNER MEMBRANE (PROBABLE).				
CC	-1- SIMILARITY: BELONGS TO THE EXEM/PULM/OUTW/XCPZ FAMILY.				
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
DR	EMBL: X56183; G45443; -				
DR	PIR: S12356; S12356.				
DR	PIR: S21639; S21639.				
FT	TRANSFORM: TRANSFORMER: INNER MEMBRANE.				
KW	TRANSMEM 33				
FT	POTENTIAL. 52				
SQ	SEQUENCE 174 AA; 19265 MW; 2160BB8F CRC32;				
Query Match	16.4%;	Score 98;	DB 1;	Length 174;	
Best Local Similarity	30.4%;	Pred. No. 1.00e-01;			
Matches	17;	Conservative	17;	Mismatches 21;	Indels 1;
DB	98	ALOGMTASAAAGLSVERIDNEGGAVALQAPFAKILPMLFOLNGQGVYA	153		
OY	20	AQLRDLTSSSDLSMTIELLEKDM-AFOEALDPFPDGGSFRAOELLDGQOAS	74		

```

RESULT 2
ID Y021-STNY3 STANDARD: PRT: 277 AA.
AC Q05682;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE PUTATIVE PROTEASE SLR0021 (EC 3.4.-.-).
GN SLR0021.
OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC BACTERIA: CYANOBACTERIA: CHROCOCCALES: SYNECHOCYSTIS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96127529.
RA KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
RA SUGIURA M., TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome."
RL DNA RES. 2:153-166(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY U7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: D64000; GI001581;
DR PFWMI: PF01343; Peptidase U7; 1.
DR HYPOTHETICAL PROTEIN; HYDROLASE; PROTEASE.
KW SEQUENCE 277 AA: 30274 MW: E996320D CRC32:
SQ
Query Match 15.7%: Score 94; DB 1: Length 277;
Best Local Similarity 39.4%: Pred. No. 3.43e-01;
Matches 13; Conservative 9; Mismatches 10; Indels 1; Gaps 1;
Db 120 LLEKGVSEK-VIKSPYKDIISFRELPEQ 151
39 LLEKDMAFOELDPGDPGSPFAOELDDQ 71
Oy
RESULT 3
ID YBBB-ECOLI STANDARD: PRT: 364 AA.
AC P33667; P77709;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 41.1 KD PROTEIN IN RHD-GCL INTERGENIC REGION.
GN YBBB.
OS ESCHERICHIA COLI.
OC BACTERIA: PROTEOBACTERIA: GAMMA SUBDIVISION: ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE: 9742661.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.R., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RT SCIENCE 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX ROBERTS D., ALLEN E., ARAUJO R., APARICIO A., CHUNG E., DAVIS K.,
RA DUNCAN M., FEDERSPIEL N., HYMAN R., KALMAN S., KOMP C., KURDI O.,
RA LEW H., LIN D., NAMATH A., OEFNER P., SCHRAMM S., DAVIS R.W.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE OF 268-364 FROM N.A.

```

```

RC STRAIN-K12;
RX MEDLINE: 92115567.
RA SADOSKY A.B., GRAY J.A., HILL C.W.;
RT "The Rhd-E subfamily of Escherichia coli K-12."
RL NUCLEIC ACIDS RES. 19:7177-7183(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: AE000156; G1786712;
DR EMBL: U82664; G1773184;
DR EMBL: L19084; G304940;
DR ECOGENE: EC11768; YBBB.
KW HYPOTHETICAL PROTEIN; ATP-BINDING.
FT NP_BIND 149 156 ATP (POTENTIAL).
SQ SEQUENCE 364 AA: 41110 MW: DA776C59 CRC32:
Query Match 15.7%: Score 94; DB 1: Length 364;
Best Local Similarity 39.4%: Pred. No. 3.43e-01;
Matches 13; Conservative 10; Mismatches 8; Indels 2; Gaps 2;
Db 298 YNELAARLDAALTLQTGTGTHLAVPLLE 330
11 FGLPDLHQLR-DLTSSSD-ELSWITELLE 41
Oy
RESULT 4
ID SYAC-YEAST STANDARD: PRT: 958 AA.
AC P40825;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ALANYL-TRNA SYNTHETASE, CYTOPLASMIC (EC 6.1.1.7) (ALANINE--TRNA
DE LIGASE) (ALARS).
GN ALA1 OR YOR335C.
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA: FUNGI: ASCOMYCOTA: HEMIASCOMYCETES: SACCAROMYCETALES;
OC SACCAROMYCETACEAE; SACCAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RA RUPMASTER T.L., SCHIMMEL P.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 9705186.
RA PARLE-MCDERMOTT A.G., HAND N.J., GOULDING S.G., WOLFE K.H.;
RT "Sequence of 29 kb around the PD10 locus on the right arm of
RT Saccharomyces cerevisiae chromosome XV: similarity to part of
RT chromosome I."
RL YEAST 12:999-1004(1996).
CC -1- CATALYTIC ACTIVITY: ATP + L-ALANINE + TRNA(ALA) -> AMP +
CC PYROPHOSPHATE + L-ALANYL-TRNA(ALA).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: U18672; G609338;
DR EMBL: Z49821; E218690;
DR EMBL: Z75243; E252159;
DR SGD: L0002757; ALA1.
DR PROSITE: PS00179; AA-TRNA_LIGASE_II_1; FALSE_NEG.

```

```

DR PROSITE: PS00339: AA TRNA LIGASE. IL 2: 1.
KW AMINOACYL-TRNA SYNTHETASE; PROTEIN BIOSYNTHESIS; LIGASE; ATP-BINDING.
FT CONFLICT 136 136 R -> S (IN REF. 1).
FT CONFLICT 465 467 KDQ -> RTR (IN REF. 1).
FT CONFLICT 840 841 FE -> LQ (IN REF. 1).
SO SEQUENCE 958 AA: 107277 MW: AE929944 CRC32:

Query Match 15.2%; Score 91; DB 1; Length 958;
Best Local Similarity 25.0%; Pred. No. 8.45e-01;
Matches 11; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

Db 351 FSTAPTLISOVODIFPELAKDPAFLPEILDEEASFKTKDRG 394
11 FGLGDQLHQLRDLTSSSSDELSWITELLEKDMAFQDALDPG 54

RESULT 5 STANDARD; PRT: 2870 AA.
AC P20471;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DE 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE 319 KD PROTEIN.
GN NDVB.
OS RHIZOBIUM MELLIOTI.
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;
OC RHIZOBIACEAE; SINORHIZOBIUM.
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE: 90153914.
RT TELPI L., DYLAN T., DITTA G.S., HELINSKI D.R., STANFIELD S.W.;
RT "The ndvb locus of Rhizobium melliotti encodes a 319-kDa protein
RT involved in the production of beta-(1->2)-glucan."
RL J. BIOL. CHEM. 265:2843-2851(1990).
CC -1- FUNCTION: INVOLVED IN THE PRODUCTION OF BETA-(1,2)-GLUCAN.
CC IT IS INVOLVED NOT ONLY IN INVASION BUT ALSO IN BACTEROID
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -1- SIMILARITY: TO A. TUMEFACIENS CHVB.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J05219; G152271; -
DR PIR: A35548; A35548.
KW TRANSMEMBRANE.
FT TRANSMEM 448 470 POTENTIAL.
FT TRANSMEM 868 889 POTENTIAL.
FT TRANSMEM 996 1017 POTENTIAL.
SO SEQUENCE 2870 AA: 319777 MW: 36CDCC2DF CRC32:

Query Match 15.2%; Score 91; DB 1; Length 2870;
Best Local Similarity 42.9%; Pred. No. 8.45e-01;
Matches 15; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

Db 275 QLYRMADGSSGCAVAMIERLERKGTVEEAL 309
17 QLHAQLRDLTSSSSDELSWITELLEKDMAFQDAL 51

RESULT 6 STANDARD; PRT: 395 AA.
AC P06726;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE APOLIPOPROTEIN A-IV PRECURSOR (APO-AIV).
DE APOA4.
GN APOA4.

OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIODONGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE: 87089722.
RA WILLIAMS S.C., BROCKHEIMER S.M., LUSIS A.J., LEBROUF R.C.,
RA KINNBURGH A.J.;
RT "Mouse apolipoprotein A-IV gene: nucleotide sequence and induction by
RT a high-lipid diet."
RT MOL. CELL. BIOL. 6:3807-3814(1986).
RN [2]
RX REVISIONS.
RA KINNBURGH A.J.;
RL SUBMITTED (DEC-1986) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J, AND 129/J;
RX MEDLINE: 91286309.
RA REUE K., LEBTE T.H.;
RT "Genetic variation in mouse apolipoprotein A-IV due to insertion and
RT deletion in a region of tandem repeats."
RL J. BIOL. CHEM. 266:12715-12721(1991).
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
CC COMPONENT OF HDL AND CHYLOMICRONS.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: SECRETED IN PLASMA.
CC -1- POLYMORPHISM: THERE IS A POLYMORPHISM WITHIN A SERIES OF IMPERFECT
CC REPEATS ENCODING THE SEQUENCE E-Q-[AV]-Q. INSERTIONS OR DELETIONS
CC OF 12 NUCLEOTIDES HAVE GIVEN RISE TO THREE FORMS CHARACTERIZED BY
CC THREE (129), FOUR (C57BL/6), OR FIVE (M.CASTANEUS) COPIES OF THE
CC REPEAT UNIT.
CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO. A AND B. RELATED 11-
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M13966; G387102; -
DR EMBL: M64249; G191887; -
DR EMBL: M64248; G191885; -
DR PIR: A25281; A25281.
KW PLASMA; LIPID TRANSPORT; HDL; CHYLOMICRON; REPEAT; SIGNAL;
KW POLYMORPHISM.
FT SIGNAL 1 20
FT CHAIN 21 395
FT DOMAIN 33 330
FT REPEAT 33 81
FT REPEAT 60 84
FT REPEAT 82 103
FT REPEAT 115 136
FT REPEAT 137 158
FT REPEAT 159 180
FT REPEAT 181 202
FT REPEAT 203 224
FT REPEAT 225 246
FT REPEAT 247 268
FT REPEAT 269 286
FT REPEAT 287 308
FT REPEAT 309 330

```

FT DOMAIN 366 389 GTU/GLN-RICH
 FT VARIANT 382 395 MISSING (IN SOME STRAINS).
 FT CONFLICT 15 15 MISSING (IN REF. 1).
 FT CONFLICT 63 63 Q -> K (IN REF. 1).
 FT CONFLICT 207 207 E -> R (IN REF. 1).
 FT CONFLICT 288 288 S -> A (IN REF. 1).
 FT CONFLICT 294 295 RO -> KA (IN REF. 1).
 FT CONFLICT 315 316 NK -> CG (IN REF. 1).
 FT CONFLICT 395 AA: 45029 MM: 3B5145BD CRC32:
 SQ SEQUENCE

Query Match 14.9%; Score 89; DB 1; Length 395;
 Best Local Similarity 31.5%; Pred. No. 1,52e+00;
 Matches 23; Conservative 20; Mismatches 27; Indels 3; Gaps 3;

Db 322 LEQFROQLPNSGEVSHSLFLEKSLREKVNSTWSTLEKKGSPDQAL-PLP-EQAOEQ 379
 4 LELRLVFPGLGQDLHAQRLDLSSSSDELSTWIELLEKDGMAFQ-EALDPGPFDOGSF 62

Db 380 AOEAOEQVQPK 392
 63 AOELEDDGOASP 75

RESULT 7 STANDARD: PRT: 916 AA.
 ID SECA TREPA
 AC 083394;
 DT 15-DEC-1998 (REL. 37, CREATED)
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE PREPROTEIN TRANSLOCASE SECA SUBUNIT.
 OS SECA OR TP0379.
 OS TREPONEMA PALLIDUM.
 OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; TREPONEMA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-NICHOLS;
 RX MEDLINE: 98332770.
 RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
 RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
 RA SODERGRAN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
 RA KHALAF H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.,
 RA MCDONALD L., ARTICHON P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
 RA HATCH B., HORST K., ROBERTS K., MATTHEY L., WEIDMAN J., SMITH H.O.,
 RA VENTER J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis
 spirochete.";
 RT SCIENCE 281:375-388(1998).
 CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH THE SECA/SPCE
 SUBUNITS. SECA HAS A CENTRAL ROLE IN COUPLING THE HYDROLYSIS OF
 ATP TO THE TRANSFER OF PRE-SECRETORY PERIPLASMIC AND OUTER
 MEMBRANE PROTEINS ACROSS THE MEMBRANE (BY SIMILARITY).
 CC -1- SUBUNIT: ONE OF SEVEN SECRETORY PROTEINS (SECA-F & SECY) THAT
 COMPREHENDS THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS
 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF PLASMA MEMBRANE
 (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SECA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE001217; G3322659; -.
 DR TIGR: TP0379; -.
 KW PROTEIN TRANSPORT; ATP-BINDING; MEMBRANE; TRANSLOCATION; TRANSPORT.
 FT NP_BIND 103 110 ATP (POTENTIAL).
 SQ SEQUENCE 916 AA: 103779 MM: 84D5AFAC CRC32:
 Query Match. 14.9%; Score 89; DB 1; Length 916;

Best Local Similarity 34.2%; Pred. No. 1,52e+00;
 Matches 13; Conservative 11; Mismatches 12; Indels 2; Gaps 2;

Db 7 LRLIFGSOHERDLKLNPLLNVAHNAOE-SWVLPLOESE 43
 7 LRLVGPGLGQDLHAQRLDLSSSSDELSTWIELLEKDG 43

RESULT 8 STANDARD: PRT: 952 AA.
 ID IF41 YEAST
 AC P39935;
 DT 01-FEB-1995 (REL. 31, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE EUKARYOTIC INITIATION FACTOR 4F SUBUNIT P150 (EIF-4F) (MRNA CAP-
 BINDING PROTEIN COMPLEX SUBUNIT P150).
 GN IF4631 OR YGRL62W.
 OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
 OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 93330281.
 RA GOYER C., ALTMANN M., LEE H.S., BLANC A., DESHMUKH M.,
 RA WOOLFORD J.L., TRACHSEL H., SONENBERG N.;
 RT "IF4631 and IF4632: two yeast genes encoding the
 high-molecular-weight subunits of the cap-binding protein complex
 (eukaryotic initiation factor 4F) contain an RNA recognition
 motif-like sequence and carry out an essential function.";
 RT MOL. CELL. BIOL. 13:4860-4874(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE: 97435481.
 RA RIEGER M., BRUECKNER M., SCHAEFER M., MUELLER-AUER S.;
 RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
 chromosome VII".
 RT YEAST 13:1077-1090(1997).
 CC -1- FUNCTION: INTERACTS WITH THE MRNA CAP STRUCTURE, M7GPPX. THIS
 INTERACTION IS REQUIRED FOR EFFICIENT RIBOSOME BINDING TO THE
 MRNA. IF4631 IS PROBABLY ESSENTIAL WHEN IF4632 IS MISSING.
 CC -1- SUBUNIT: THE CAP-BINDING PROTEIN COMPLEX IS COMPOSED OF AT LEAST
 TWO PROTEINS, A 24 KD (IF45) AND A 150-200 KD SUBUNIT (IF4631).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNP).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L16923; G295675; -.
 DR EMBL: Z72847; E243549; -.
 DR PIR: A48086; A48086.
 DR SGD: L0002309; TIF4631.
 DR PROSITE: PS00030; RNP_1; FALSE_NEG.
 KW INITIATION FACTOR; PROTEIN BIOSYNTHESIS; RNA-BINDING;
 KM MULTIGENE FAMILY.
 FT DOMAIN 173 200 PRO/SER/THR-RICH.
 FT DOMAIN 375 384 ALA/GLU-RICH.
 FT DOMAIN 488 553 ARG/SER-RICH.
 FT DOMAIN 873 899 ARG/SER-RICH.
 FT DOMAIN 869 872 POLY-GLU.
 FT DOMAIN 647 652 RNA-BINDING (RNP2) (BY SIMILARITY).
 FT DOMAIN 742 749 RNA-BINDING (RNP1) (BY SIMILARITY).
 FT CONFLICT 7 7 H -> Q (IN REF. 1).
 FT CONFLICT 37 37 T -> N (IN REF. 1).
 FT CONFLICT 110 110 Q -> K (IN REF. 1).
 FT CONFLICT 207 207 R -> K (IN REF. 1).
 FT CONFLICT 361 361 D -> E (IN REF. 1).
 SQ SEQUENCE 952 AA: 107101 MM: DEB23EB9 CRC32:

Query Match 14.9% Score 89; DB 1; Length 952;
Best Local Similarity 32.7%; Pred. No. 1.52e+00;
Matches 16; Conservative 15; Mismatches 15; Indels 3; Gaps 3;

Db 772 MNDLTPSEETLESVELLWVGEOPFDSFTGATLEGSLDLSL 820
22 LNDLTSSSDSDE-LSWITTELEKGMFQ-EALDPGPF-DGSPFAOELL 67

RESULT 9
ID YGFB.ECOLI STANDARD; PRT: 194 AA.
AC P25533;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHEICAL 21.5 KD PROTEIN IN PEPP-SSR INTERGENIC REGION (ORF194)
DE (F194).
GN YGFB.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-HB101;
MEDLINE: 930534351.
RA NAKAHIGASHI K., MIYAMOTO K., NISHIMURA K., INOKUCHI H.;
RT Isolation and characterization of a light-sensitive mutant of
RT Escherichia coli K-12 with a mutation in a gene that is required for
RT the biosynthesis of ubiquinone."
RL J. BACTERIOL. 174:7352-7359(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
MEDLINE: 97426617.
RA BLATTNER F.R., PLUNKETT G., III, BLOCH C.A., PERA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RT SCIENCE 277:1453-1474(1997).
CC -1- SIMILARITY: STRONG, TO H.INFLUENZA H10817.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D90281; G216626;
DR EMBL: U28377; G882439;
DR EMBL: AE000374; G1789276;
DR PIR: J00842; J00842;
DR PIR: A47020; A47020;
DR ECOGENE: EG11323; YGFB.
KW HYPOTHEICAL PROTEIN.
SO SEQUENCE 194 AA; 21474 MW; 104C05A8 CRC32;

Query Match 14.7% Score 88; DB 1; Length 194;
Best Local Similarity 28.6%; Pred. No. 2.04e+00;
Matches 16; Conservative 16; Mismatches 22; Indels 2; Gaps 1;

Db 30 EMHGLTSGMTCGNDSSWPLJLHDJLTNEGMAEALRKMHSATSALDDG 85
17 QLMQAQRLDTLSSSDLSLWITEL-LEKDMARQALDPCGPFDOGSPFAOELLDDG 70

RESULT 10
ID RPEP.ALCEU STANDARD; PRT: 241 AA.
AC 004539;
DT 01-FEB-1995 (REL. 31, CREATED)

DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE RIBULOSE-5-PHOSPHATE 3-EPIMERASE, PLASMID (EC 5.1.3.1) (PENTOSE-5-
DE PHOSPHATE 3-EPIMERASE) (PPE) (R5P3E).
GN CBEP OR CPEX.
OS ALCALIGENES EUTROPHUS.
OC PLASMID MEGAPLASMID PHG1.
OC BACTERIA; PROTEOBACTERIA; BETA SUBDIVISION; BURKHOLDERIA GROUP;
OC RALSTONIA.

RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-H16 / ATCC 17699;
MEDLINE: 93053439.
RA KUSTAN B., YOO J.-G., BEDNARSKI R., BOWEN B.;
RT "The Calvin cycle enzyme pentose-5-phosphate 3-epimerase is encoded
RT within the cfx operons of the chemolithotroph Alcaligenes eutrophus."
RL J. BACTERIOL. 174:7337-7344(1992).
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 5-PHOSPHATE -> D-XYLULOSE 5-
CC PHOSPHATE.
CC -1- PATHWAY: CALVIN CYCLE.
CC -1- SIMILARITY: BELONGS TO THE RIBULOSE-PHOSPHATE 3-EPIMERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M64172; G150678;
DR PROSITE: PS01085; RIBUL_P_3-EPIMER_1; 1.
DR PROSITE: PS01086; RIBUL_P_3-EPIMER_2; 1.
DR PFAM: PF00834; Ribul_P_3-epim; 1.
KW ISOMERASE; CARBOHYDRATE METABOLISM; CALVIN CYCLE; PLASMID.
SO SEQUENCE 241 AA; 25594 MW; 4F4B4BCD CRC32;

Query Match 14.7% Score 88; DB 1; Length 241;
Best Local Similarity 30.6%; Pred. No. 2.04e+00;
Matches 15; Conservative 13; Mismatches 19; Indels 2; Gaps 2;

Db 126 LVNRPATPSLMDHLDKLDVLMNVNPG-FG-GQAPFVLDKVRQA 172
25 LNSSSDLSLWITTELEKGMFQ-EALDPGPFDOGSPFAOELLDDGQA 73

RESULT 11
ID SPCA.HUMAN STANDARD; PRT: 2418 AA.
AC P02549; Q15514;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE SPECTRIN ALPHA CHAIN, ERYTHROCYTE.
GN SPTAL OR SPTA.
OS HOWO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 90170949.
RA SAHR K.E., LAURITA P., KOTULA L., SCARPA A.L., COUPAL E., LETO T.L.,
RA LINNENBACH A.J., WINKELMANN J.C., SPEICHER D.W., MARCHESI V.T.,
RA CURTIS P.J., FORGET B.G.;
RT "The complete cDNA and polypeptide sequences of human erythroid
RT alpha-spectrin."
RL J. BIOL. CHEM. 265:4434-4443(1990).
RN [2]
RP SEQUENCE OF 1-533 FROM N.A.
RX MEDLINE: 90009318.
RA SAHR K.E., TOBE T., SCARPA A., LAUGHINGHOUSE K., MARCHESI S.L.,
RA AGRE P., LINNENBACH A.J., MARCHESI V.T., FORGET B.G.;
RT "Sequence and exon-intron organization of the DNA encoding the alpha
RT I domain of human spectrin. Application to the study of mutations

RT causing hereditary elliptocytosis.";
 RL J. CLIN. INVEST. 84:1243-1252(1989).
 [13]
 RP SEQUENCE OF 7-601.
 RX MEDLINE: 84087888.
 RA SPEICHER D.W., DAVIS G., MARCHESI V.T.;
 RT "Structure of human erythrocyte spectrin. II. The sequence of the
 RL alpha-I domain.";
 RL J. BIOL. CHEM. 258:14938-14947(1983).
 [14]
 RP SEQUENCE OF 7-125.
 RX MEDLINE: 84087887.
 RA SPEICHER D.W., DAVIS G., YURCHENCO P.D., MARCHESI V.T.;
 RT "Structure of human erythrocyte spectrin. I. Isolation of the alpha-I
 RL domain and its cyanogen bromide peptides.";
 RL J. BIOL. CHEM. 258:14931-14937(1983).
 [15]
 RP SEQUENCE OF 320-450 FROM N.A.
 RX MEDLINE: 86205962.
 RA LINNENBACH A.J., SPEICHER D.W., MARCHESI V.T., FORGET B.G.;
 RT "Cloning of a portion of the chromosomal gene for human erythrocyte
 RL alpha-spectrin by using a synthetic gene fragment.";
 RL PROC. NATL. ACAD. SCI. U.S.A. 83:2397-2401(1986).
 [16]
 RP SEQUENCE OF 7-533 FROM N.A., AND VARIANT PRO-260; PRO-261 AND PRO-471.
 RX MEDLINE: 90009318.
 RA SAHR K.E., TOBE T., SCARPA A.L., LAUGHINGHOUSE K., MARCHESI S.L.,
 AGRE P., LINNENBACH A.J., MARCHESI V.T., FORGET B.G.;
 RT "Sequence and exon-intron organization of the DNA encoding the alpha
 RL I domain of human spectrin. Application to the study of mutations
 RT causing hereditary elliptocytosis.";
 RL J. CLIN. INVEST. 84:1243-1252(1989).
 [17]
 RP PARTIAL SEQUENCE.
 RX MEDLINE: 84295638.
 RA SPEICHER D.W., MARCHESI V.T.;
 RT "Erythrocyte spectrin is comprised of many homologous triple helical
 RL segments.";
 RL NATURE 311:177-180(1984).
 [18]
 RP IDENTIFICATION OF PROBABLE FRAMESHIFT IN 2407-2418.
 RA GIBSON T.J.;
 RL UNPUBLISHED OBSERVATIONS (MAR-1995).
 [19]
 RP REVIEW ON VARIANTS.
 RX MEDLINE: 97001215.
 RA MAILLET P., ALLOISIO N., MORLE L., DELAUNAY J.;
 RT "Spectrin mutations in hereditary elliptocytosis and hereditary
 RL spherocytosis.";
 RL HUM. MUTAT. 8:97-107(1996).
 [10]
 RP VARIANT SER-24.
 RX MEDLINE: 94289716.
 RA PARQUET N., DEVAUX I., BOULANGER L., GALAND C., BOIVIN P.,
 LECOMTE M.-C., DHERMY D., GARBARZ M.;
 RT "Identification of three novel spectrin alpha I/74 mutations in
 RL hereditary elliptocytosis: further support for a triple-stranded
 RL folding unit model of the spectrin heterodimer contact site.";
 RL BLOOD 84:303-308(1994).
 [11]
 RP VARIANTS CYS-28; HIS-28; LEU-28 AND SER-28.
 RX MEDLINE: 91358728.
 RA COETZER T.L., SAHR K., PRCHAL J., BLACKLOCK H., PETERSON L., KOLER R.,
 DOYLE J., MANASTER J., PALEK J.;
 RT "Four different mutations in codon 28 of alpha spectrin are
 RL associated with structurally and functionally abnormal spectrin alpha
 RL I/74 in hereditary elliptocytosis.";
 RL J. CLIN. INVEST. 88:743-749(1991).
 [12]
 RP VARIANTS SER-28 AND ARG-48.
 RX MEDLINE: 91346849.
 RA FLOYD P.B., GALLAGHER P.G., VALENTINO L.A., DAVIS M., MARCHESI S.L.,
 FORGET B.G.;

RT "heterogeneity of the molecular basis of hereditary
 RT pyropoikilocytosis and hereditary elliptocytosis associated with
 RT increased levels of the spectrin alpha I/74-kilodalton tryptic
 RL peptide.";
 RL BLOOD 78:1364-1372(1991).
 [13]
 RP VARIANT SER-45.
 RX MEDLINE: 89323468.
 RA LECOMTE M.-C., GARBARZ M., GRANDCHAMP B., FEO C., GAUTERO H.,
 DEVAUX I., BOURNIER O., GALAND C., D'AURIOL L., GALIBERT F.,
 SAHR K.E., FORGET B.G., BOIVIN P., DHERMY D.;
 RT "Sp alpha I/78: a mutation of the alpha I spectrin domain in a white
 RL kindred with HE and HRP phenotypes.";
 RL BLOOD 74:1126-1133(1989).
 [14]
 RP VARIANT PRO-207.
 RX MEDLINE: 92176375.
 RA GALLAGHER P.G., TSE W.T., COETZER T., LECOMTE M.-C., GARBARZ M.,
 ZARKOWSKY H.S., BARUCHEL A., BALLAS S.K., DHERMY D., PALEK J.,
 FORGET B.G.;
 RT "A common type of the spectrin alpha I 46-50a-KD peptide abnormality
 RL in hereditary elliptocytosis and pyropoikilocytosis is associated
 RL with a mutation distant from the proteolytic cleavage site. Evidence
 RT for the functional importance of the triple helical model of
 RL spectrin.";
 RL J. CLIN. INVEST. 89:892-898(1992).
 [15]
 RP VARIANT VAL-1857.
 RX MEDLINE: 93253053.
 RA WILMOTTE R., MARECHAL J., MORLE L., BAKLOUTI F., PHILIPPE N.,
 KASTALLY R., KOTUA L., DELAUNAY J., ALLOISIO N.;
 RT "Low expression allele alpha LEIX of red cell spectrin is associated
 RL with mutations in exon 40 (alpha V/41 polymorphism) and intron 45 and
 RL with partial skipping of exon 46.";
 RL J. CLIN. INVEST. 91:2091-2096(1993).
 [16]
 RP VARIANT BARCELONA.
 RX MEDLINE: 93572367.
 RA DELLA VENEZIA N., ALLOISIO N., FORISSTER A., DENOROV L., AYMERICH M.,
 VIVES-CORRONS J.L., BESALDUCH J., BESSON I., DELAUNAY J.;
 RT "Elliptocytosis associated with the alpha 469 His-->Pro
 RL mutation in spectrin Barcelona (alpha I/50-46b).";
 RL BLOOD 82:1661-1665(1993).
 [17]
 RP VARIANT CAGLIARI.
 RX MEDLINE: 94043025.
 RA SAHR K.E., COETZER T.L., MOY L.S., DERICK L.H., CHISHTI A.H.,
 JAROLIM P., LORENZO F., MIRAGLIA DEL GIUDICE E., IOLASCON A.,
 GALLANELLO R.;
 RT "Spectrin Cagliari: an Ala-->Gly substitution in helix 1 of beta
 RL spectrin repeat 17 that severely disrupts the structure and self-
 RL association of the erythrocyte spectrin heterodimer.";
 RL J. BIOL. CHEM. 268:22656-22662(1993).
 [18]
 RP VARIANTS CUIOZ AND LYON.
 RX MEDLINE: 90347052.
 RA MORLE L., ROUX A.-F., ALLOISIO N., POTHIER B., STARCK J., DENOROV J.,
 MORLE F., RUDIGLOZ R.C., FORGET B.G., DELAUNAY J., GOERT J.;
 RT "Two elliptocytogenic alpha I/74 variants of the spectrin alpha I
 RL domain. Spectrin Cuioz (GGT-->GTT; alpha I 40 Gly-->Val) and
 RL spectrin Lyon (CTT-->TTT; alpha I 43 Leu-->Phe).";
 RL J. CLIN. INVEST. 86:548-554(1990).
 [19]
 RP VARIANT JENDOUBA.
 RX MEDLINE: 92345619.
 RA ALLOISIO N., WILMOTTE R., MORLE L., BAKLOUTI F., MARECHAL J.,
 DECUZEAU M.-T., DENOROV L., FEO C., FORGET B.G., KASTALLY R.,
 DELAUNAY J.;
 RT "Spectrin Jendouba: an alpha II/31 spectrin variant that is
 RL associated with elliptocytosis and carries a mutation distant from
 RL the dimer self-association site.";
 RL BLOOD 80:809-815(1992).
 [20]

RP VARIANT TUNIS.
 RA MEDLINE: 89323436.
 RA MORIE L., MORIE F., ROUX A.F., GODET J., FORGET B.G., DENOROV L.,
 RA GABBAR M., DHERY D., KASTALLY R., DELAUNAY J.,
 RT "Spectrin Tunis (SP alpha I/78), an elliptocytogenic variant, is due
 RT to the CGG->TGG codon change (Arg->Tyr) at position 35 of the
 RT alpha I domain.".
 RL BLOOD 74:828-832(1989).
 RN [12]
 RN VARIANT GENOVA.
 RX MEDLINE: 94250920.
 RA PERROTTA S., DEL GUIDICE E.M., ALLOISIO N., SCIARRETTA G., PINTO L.,
 RT "Mild elliptocytosis associated with the alpha 34 Arg->Tyr mutation
 RT in spectrin genova (alpha I/74).".
 RL BLOOD 83:3346-3349(1994).
 RN [12]
 RN VARIANT ANASTASIA.
 RX MEDLINE: 95290423.
 RA PERROTTA S., IOLASCON A., DE ANGELIS F., PAGANO L., COLONNA G.,
 RA CURTILLO S., DEL GUIDICE E.M.,
 RT "Spectrin Anastasia (alpha I/78): a new spectrin variant (alpha 45
 RT Arg->Thr) with moderate elliptocytogenic potential.".
 RL BR. J. HAEMATOL. 89:933-936(1995).
 CC -I- FUNCTION: SPECTRIN IS THE MAJOR CONSTITUENT OF THE CYTOSKELETAL
 CC NETWORK UNDERLYING THE ERYTHROCYTE PLASMA MEMBRANE. IT ASSOCIATES
 CC WITH BAND 4.1 AND ACTIN TO FORM THE CYTOSKELETAL SUPERSTRUCTURE OF
 CC THE ERYTHROCYTE PLASMA MEMBRANE.
 ...
 Note: remainder of annotations omitted.

Query Match 14.7%; Score 88; DB 1; Length 2418;
 Best Local Similarity 31.3%; Pred. No. 2.04e+00;
 Matches 15; Conservative 15; Mismatches 15; Indels 3; Gaps 3;
 DB 1318 AEDLGIELLEERH-QEHRADMEAPTFQALDESAEIDSGHHASP 1364
 Oy 30 SDELSWTELEKDMARFOEALDP-GP-FDQGSFPAQELHDDGQASP 75

RESULT 12
 ID YDJA_ECOLI STANDARD: PRT: 183 AA.
 AC P24250:
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HYPOTHEICAL 20.1 KD PROTEIN IN SELD-SPA INTERGENIC REGION (ORF183).
 GN YDJA.
 OS ESCHERICHIA COLI.
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
 OC ESCHERICHIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 91317715.
 RA SAMERS G., HEIDER J., ZEHLEFEN E., BOECK A.,
 RT "Expression and operon structure of the sel genes of Escherichia coli
 RT and identification of a third selenium-containing formate
 RT dehydrogenase isoenzyme.".
 RL J. BACTERIOL. 173:4983-4993(1991).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RA MEDLINE: 97426617.
 RA BLATTNER F.R., PUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
 RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
 RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GODDEN M.A., ROSE D.J.,
 RA MAU B., SHAO Y.,
 RT "The complete genome sequence of Escherichia coli K-12.".
 RL SCIENCE 277:1453-1474(1997).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12.
 RX MEDLINE: 97251357.

RA AIBA H., BABI T., FUJITA K., HAYASHI K., INADA T., ISONO K.,
 RA ITOH T., KASAI H., KASHIMOTO K., KIMURA S., KITARAWA M.,
 RA KITAGAWA M., MAKINO K., MIKI T., MIZOBUCHI K., MORI H., MORI T.,
 RA MOTOMURA K., NAKADE S., NAKAMURA Y., NASHIMOTO H., NISHIO Y.,
 RA OSHIMA T., SAITO N., SAMPEI G., SEKI Y., SIVASUNARAM S.,
 RA TAGAMI H., TAKEDA J., TAKEUCHI K., TAKEUCHI Y., WADA C.,
 RA YAMAMOTO Y., HORIUCHI T.,
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map.".
 RL DNA RES. 3:363-377(1996).
 CC -I- SIMILARITY: STRONG. TO H. INFLUENZAE H11542.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M68961; G147021; -
 DR EMBL: AE000271; G1788063; -
 DR EMBL: D90820; G1742876; -
 DR PIR: A40360; A40360.
 DR ECGENE: EG1134; YDJA.
 DR PFAM: PF00881; Nitroreductase; 1.
 KW HYPOTHEICAL PROTEIN.
 SQ SEQUENCE 183 AA; 20059 MW; 6909542D CRC32;

Query Match 14.4%; Score 86; DB 1; Length 183;
 Best Local Similarity 26.9%; Pred. No. 3.63e+00;
 Matches 14; Conservative 15; Mismatches 20; Indels 3; Gaps 3;
 DB 20 APTGEOLNIRACGRAPDRHSMOPHFVEYEGEREFSALFCGATAGS 71
 Oy 12 GPLDQLHQLARD-LTSSSDBL-SWTELEKMGMA-FQALDPGPDDGS 60

RESULT 13
 ID GRP3_ARTSA STANDARD: PRT: 308 AA.
 AC P13230:
 DT 01-JAN-1990 (REL. 13, CREATED)
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE GLYCINE-RICH PROTEIN GRP3.
 DE ARTEMISA SALINA (BRINE SHRIMP).
 OS EUKARYOTA; METAZOA; ARTHROPODA; CRUSTACEA; BRANCHIOPODA; ANOSTRACA;
 OC ARTEMIDIACE; ARTEMIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 88007550.
 RA CRUZ-ALVAREZ M., PELLICER A.,
 RT "Cloning of a full-length complementary DNA for an Artemia salina
 RT glycine-rich protein. Structural relationship with RNA binding
 RT proteins.".
 RL J. BIOL. CHEM. 262:13377-13380(1987).
 CC -I- PIM: THE ARGININES IN THE GLY-RICH DOMAIN MIGHT BE METHYLATED.
 CC -I- SIMILARITY: TO HD40, THE MAJOR PROTEIN COMPONENT OF ARTEMIA
 CC HEMERGENEUS NUCLEAR RIBONUCLEOPROTEIN PARTICLES, AND
 CC STRUCTURALLY TO OTHER NUCLEAR RNA BINDING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J03453; G161174; -
 DR PIR: A29379; A29379.
 KW NUCLEAR PROTEIN; RIBONUCLEOPROTEIN; METHYLATION.
 FT DOMAIN 83 118 KH.

RP VARIANT TUNIS.
 RX MEDLINE: 89323436.
 RA MOBLE L., MOBLE F., ROUX A.F., GODET J., FORGET B.G., DENOROV L.,
 RA GABBAZ M., DHERRY D., KASTALLY R., DELAUNAY J.,
 RT "Spectrin Tunis (Sp alpha I/78), an elliptocytogenic variant, is due
 to the CGG->TGC codon change (Arg->Tyr) at position 35 of the
 alpha I domain."
 RL BLOOD 74:828-832(1989).
 RN [21]
 RP VARIANT GENOVA.
 RX MEDLINE: 94250920.
 RA PEROTTA S., DEL GUIDICE E.M., ALLOISIO N., SCIARATTA G., PINTO L.,
 RA DELAUNAY J., CUTILLO S., IOLASCON A.,
 RT "Mild elliptocytosis associated with the alpha 34 Arg->Tyr mutation
 in spectrin Genova (alpha I/74)."
 RL BLOOD 83:3346-3349(1994).
 RN [22]
 RP VARIANT ANASTASIA.
 RX MEDLINE: 95290423.
 RA PEROTTA S., IOLASCON A., DE ANGELIS F., PAGANO L., COLONNA G.,
 RA CUTILLO S., DEL GUIDICE E.M.,
 RT "Spectrin Anastasia (alpha I/78): a new spectrin variant (alpha 45
 Arg->Tyr) with moderate elliptocytogenic potential."
 RL BR. J. HAEMATOL. 89:933-936(1995).
 CC -I- FUNCTION: SPECTRIN IS THE MAJOR CONSTITUENT OF THE CYTOSKELETAL
 NETWORK UNDERLYING THE ERYTHROCYTE PLASMA MEMBRANE. IT ASSOCIATES
 WITH BAND 4.1 AND ACTIN TO FORM THE CYTOSKELETAL SUPERSTRUCTURE OF
 THE ERYTHROCYTE PLASMA MEMBRANE.
 CC
 Note: remainder of annotations omitted.

Query Match 14.7%; Score 88; DB 1; Length 2418;
 Best Local Similarity 31.3%; Pred. No. 2,04e+00;
 Matches 15; Conservative 15; Mismatches 15; Indels 3; Gaps 3;
 Db 1318 AEDLTGIEILERH-QEHRADMEAEAPTEQALDEPSAELIDSGHASP 1364
 YQ 30 SDELSWITELLEKMGMAFOEALDP-GP-PDQSPFAOEILDGQDASP 75

RESULT 12
 ID YDJA.ECOLI STANDARD; PRT: 183 AA.
 AC P24250;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HYPOTHEETICAL 20.1 KD PROTEIN IN SELD-SPPA INTERGENIC REGION (ORF183).
 GN YDJA.
 OS ESCHERICHIA COLI.
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
 CC ESCHERICHIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 91317715.
 RA SAVERS G., HEIDER J., ZEHLEIN E., BOECK A.,
 RT "Expression and operon structure of the sel genes of *Escherichia coli*
 and identification of a third selenium-containing formate
 dehydrogenase isoenzyme."
 RL J. BACTERIOL. 173:4983-4993(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-K12/MG1655;
 RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
 RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
 RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GODEN M.A., ROSE D.J.,
 RA MAU B., SHAO Y.,
 RT "The complete genome sequence of *Escherichia coli* K-12."
 RL SCIENCE 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN-K12;
 RX MEDLINE: 97251357.

RA AIBA H., BABA T., FUJITA K., HAYASHI K., INADA T., ISONO K.,
 RA ITOH T., KASHI H., KASHIMOTO K., KIMURA S., KITAKAWA M.,
 RA KITAGAWA M., MARINO K., MIKI T., MIZOBACHI K., MORI H., MORI T.,
 RA MOTOMURA K., NAKADE S., NAKAMURA Y., NASHIMOTO H., NISHIO Y.,
 RA OSHIMA T., SAITO N., SAMPEI G., SEKI Y., SVASUNPARAM S.,
 RA TAGAMI H., TAKEDA J., TAKEMOTO K., TAKEUCHI Y., WADA C.,
 RA YAMAMOTO Y., HORIUCHI T.,
 RT "A 570-kb DNA sequence of the *Escherichia coli* K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map."
 RL DNA RES. 3:363-377(1996).
 CC -I- SIMILARITY: STRONG, TO H. INFLUENZAE H11542.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC
 CC EMBL: M68961; G147021; -
 CC EMBL: A6000271; G178063; -
 CC EMBL: D90820; G1742876; -
 CC PIR: A40360; A40360.
 CC ECGENE: EG1134; YDJA.
 CC PFAM: PF00881; Nitroreductase; 1.
 CC DR
 CC KW
 CC SEQUENCE 183 AA; 20059 MW; 6909542D CRC32;

Query Match 14.4%; Score 86; DB 1; Length 183;
 Best Local Similarity 26.9%; Pred. No. 3,63e+00;
 Matches 14; Conservative 15; Mismatches 20; Indels 3; Gaps 3;
 Db 20 APTGEOLNIIIRAGMARPDHRSKMPWFVEYEGGEREFSVILGCAIAGS 71
 YQ 12 GPLDQLQAQLRD-LTSSSEDL-SWITELLEKMGMA-FOEALDPGRPDQS 60

RESULT 13
 ID GRP3.ARTSA STANDARD; PRT: 308 AA.
 AC P13230;
 DT 01-JAN-1990 (REL. 13, CREATED)
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE GLYCINE-RICH PROTEIN GRP33.
 OS ARTEMIA SALINA (BRINE SHRIMP).
 CC EUKARYOTA; METAZOA; ARTHROPODA; CRUSTACEA; BRANCHIOPODA; ANOSTRACA;
 CC ARTEMIIDAE; ARTEMIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 88007550.
 RA CRUZ-ALVAREZ M., PELICER A.,
 RT "Cloning of a full-length complementary DNA for an *Artemia salina*
 glycine-rich protein. Structural relationship with RNA binding
 proteins."
 RL J. BIOL. CHEM. 262:13377-13380(1987).
 CC -I- SIMILARITY: TO HD40, THE MAJOR PROTEIN COMPONENT OF ARTEMIA
 HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN PARTICLES, AND
 STRUCTURALLY TO OTHER NUCLEAR RNA BINDING PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J03453; G161174; -
 CC PIR: A29379; A29379.
 CC DR
 CC KW
 CC DOMAIN 83 118 KH.

